Indicio

A Beginner's Guide to Decentralized Identity and Verifiable Data

Trevor Butterworth
VP Communications & Governance

July 25, 2023

Indicio

The internet was built without a way to identify people, organizations, or things other than computers

Now, there are 5+ billion people and 20+ billion devices and tens of millions of businesses and organizations

If you are one of the 5.18 billion people online, you are likely to have a digital "identity"—or more precisely, a digital identifier. In fact, you probably have many of these digital identity-identifiers — email addresses, social media accounts, and a whole bunch of user accounts and profiles to purchase goods and access services. You may even use one of these identity-identifiers, such as a Facebook profile, to access multiple services.

We call these digital identities "centralized." This is because the personal data you provide to create an email address or a user account is stored by the entity you are opening an account with in a "central" database, and the management and verification of your identity and what you have access to all happens within that single environment.

One of the consequences of centralizing identity is that identifiers and identities are better thought of as *leased* rather than owned. You provide information about yourself, and that information is verified to a greater or lesser degree (for example, you might submit a phone number to the entity and then input a code sent to your phone by that entity to prove you are the person they are interacting with). The email provider, business, or service then stores that account information on a database under its control.

This is how digital identity and its verification have come to be managed online. "Federated" identity, where one identity account can be used to access several sites and services, is still centralized because your identity account and the personal data associated with it are held and controlled by the original identity provider.

We do things this way because the internet wasn't designed to manage human or organizational identity; it was designed to recognize only computers because that's what the network was built to connect (and for many years the number of connected computers was so low, it was possible for everyone to know who was using them). No one envisaged that the internet would evolve into the World Wide Web and connect 65 percent (and counting) of the world's population — along with billions more devices.

Centralized identity provided a way for all these people to participate online, which is an amazing achievement. But it is a far from perfect solution to the structural absence of a reliable identity layer for people and organizations and connected devices other than computers.

We must solve the underlying verification problem so that data can be trusted

The structure of centralized identity — usernames, passwords, emails — make digital identities and identifiers easy to fake or steal, whether it is through guessing a simple password (1, 2, 3, 4, 5…) or using a phishing email or robocall to get you to enter your account details. Even if you do nothing wrong, the database holding massive amounts of personal data is an attractive target for identity theft. The centralized model means that if there is a single point of failure in the defense of this database, the entire database can be compromised. The cost of this fraud is enormous in both the losses incurred and the costs of trying to keep this information secure.

Similarly, if a centralized identity system fails, you are shut out of your account — or several accounts if the identity is federated. And then, on top of all these external threats, there are internal problems: Who has access to my data and what did I agree to let them do with it by ticking the "consent" box when I opened my account?

It was once said that no one cared about data privacy given the convenience of doing everything online; but that is no longer the case. The expansive growth in data collection and surveillance has driven both public and political concern, the latter resulting in extensive data privacy laws such such as GDPR — the European Union's General Data Protection Regulation — and other global variants, all of which aim to protect people's personal data from being exploited. Storing personal data is now a compliance headache for businesses and organizations.

Finally, no one loves the friction of passwords. If they are hard to crack, they are hard to remember, and if they are easy to remember, they are easy to crack. Multifactor authentication just adds more friction when the overall course of the digital world is to make things as simple and seamless as possible.

As behavioral psychologists have noted, we are cognitive misers: we want to conserve mental bandwidth and we do so by relying on shortcuts to thinking and action. If people keep failing at online security, then maybe online security needs to be designed to address how people behave, rather than how we want them to behave?

All these problems combine to make digital interaction more complex and costly and more untrustworthy even as businesses, consumers, and the public sector all seek and embrace more digitalization. The need for billions of connected machines to have identifiers amplifies the scale of the problem; the evolution of artificial intelligence expands the scope of fakery and fraud. We must solve the underlying verification problem so that data can be trusted.

This is where decentralized identity steps in.

What is decentralized identity?

Decentralized identity addresses the following questions:

1. **Who should control my digital identity-identifiers?**
2. **Where should my personal data be stored?**
3. **How should my personal data be shared?**
4. **And how can all this be done in a way that enables a digital identity and the data associated with it to be verified as trustworthy?**

In addition, make the technology that does this easy to integrate into existing infrastructure, easy to use by everyone, privacy-preserving, frictionless, scalable, and low cost! Done. If you want to skip ahead, you can get all this in a virtual box from Google Marketplace. But if you're finding the journey illuminating, let's start with how these questions were answered.

1. Who should control my digital identity-identifier?

Answer: I should! The key technical development allowing me to do this is a decentralized identifier, aka a "DID." Think about phone numbers, emails, and websites. In a sense, they're all addresses for either contacting someone or some entity or accessing resources or services.

A DID is also a kind of address — one that you can create to represent yourself, or an organization, or a thing that's under your control.

But unlike a phone number, email, or website address, *you* fully control your DID. It doesn't require an identity provider or a certificate authority or a central registry to exist.

Hold that thought.

2. Where should my personal data be stored?

Answer: With me! If I can create a digital address for my identity, I can accept information sent to that address, such as an ID credential like a digital passport or driver's license. I can store these credentials with this information in special software on my mobile device — a digital wallet app — or in the cloud.

Again, hold that thought.

3. How should my personal data be shared?

Answer: By me choosing to share it. If I can hold my personal data, I can consent to sharing it. Of course, there are many situations where I am legally obligated to share data or prove my identity. But there are many more where that isn't the case.

1 + 2 + 3 =

You probably have many questions about how all this works, and we will get to those. First, let's summarize:

- I can create an address for my identity.
- I can store my personal data (or any other important information about an organization or a thing).
- I can control who I choose to share that data with.

What we've done is remove the need for centralized databases filled with our personal data in order to create and manage our digital identities and identifiers. This has significant implications for privacy and security. We also have control over our personal data.

That little checkbox where we tick away our personal data so advertisements can follow us around the internet (and for other, unknown purposes) no longer has the power it once had.

4. How can all this be done in a way that enables a digital identity and the data associated with it to be verified as trustworthy?

Answer: Verifiable credentials. This is where things get even more exciting and, also, complicated. First, the view from 40,000 feet: the owner of a DID can cryptographically prove that they control their DID.

How? Shortest possible version: When I create a DID, I define what the DID is an address to — me personally, the department of motor vehicles, a business, a bank. In my case, I'm the controller; in the case of a business or government agency, there may be multiple controllers.

So far, so good.

Each DID has a public cryptographic key paired with a private cryptographic key. (What are cryptographic keys? Strings of characters with algorithms—functions—that encrypt and decrypt data). When, for example, I contact my bank's DID, I use the bank's public key to encrypt my message. Only the bank has the private key to decrypt my message. *If the bank responds to my message, I know it controls the bank DID*. And because I am using my DID to connect to the bank, *the bank can also verify that I control the DID in my name, and therefore am the account owner*.

All this happens instantaneously and through a direct communications channel created between the DIDs called DIDComm. In other words, DIDs can connect and mutually authenticate each other over encrypted, one-to-one communications channels.

What this means: we now have a way to verify the source of any information.

Security bonus: because you can create any number of DIDs, each interaction with another DID can be uniquely encrypted, thereby preventing digital interactions from being correlated.

From DIDs to sharing verifiable data

So now we can both control our digital identities and mutually authenticate them before sharing data with each other; but how do I find someone, or something's DID for all this to happen? We use a verifiable data registry or distributed ledger. These are blockchains for identity and they are used for writing and verifying public DIDs. You use a distributed ledger network to write addresses and public keys and by doing so, DIDs can connect and verify each other.

As information is written to and stored by a blockchain-based ledger in a special, time-stamped way, any attempt to alter that information breaks the time-stamped chain of blocks and is detectable. Additionally, because there are multiple copies of a ledger on a network, each copy of the ledger records and distributes what is written to any one ledger. All the copies of the ledger in the network must be in agreement ("consensus") as to what is written to the ledger.

What this means: A distributed ledger network provides a way to find and authenticate the addresses for sources of information *and* prove information hasn't been tampered with.

Now that we have this architecture for creating trusted relationships, how is the actual data shared? We started out this section with the answer, "verifiable credentials." A credential is a digital schematic for assembling information. A credential for a passport contains a description of the passport's structure, the fields for specific information. A government issuing a passport credential will write this "schema" to a distributed ledger network, along with information that links this schema to the passport office.

The passport office will send a QR code to the passport applicant. This will enable the potential holder to connect over DIDComm (if the person doesn't have a digital wallet app to enable this, they will be prompted to download one).

Once a DIDComm connection has been established, the person will be offered their passport as a verifiable credential. If they accept, they will receive a credential populated with their specific data. Only the passport office (the issuer) and you, the person who has just received a verifiable credential, hold your actual passport data.

This data isn't written to the distributed ledger network. Personal data should never be written to a ledger — and this is the iron law of decentralized identity networks.

We use DIDs and DIDComm and the schema and other credential metadata on the network to do all the work of verification. This means that a verifiable credential can be verified anywhere (as long as you have the appropriate software) and at any time without checking in with the source of the data or cross checking the data with a third party.

One additional important note. Blockchain-based networks have come under scrutiny for energy consumption. This concern does not apply to blockchain-based identity networks for the reason that no "mining" is required for verification. Writing a schema to a ledger or looking up a DID consumes no more energy than a web search.

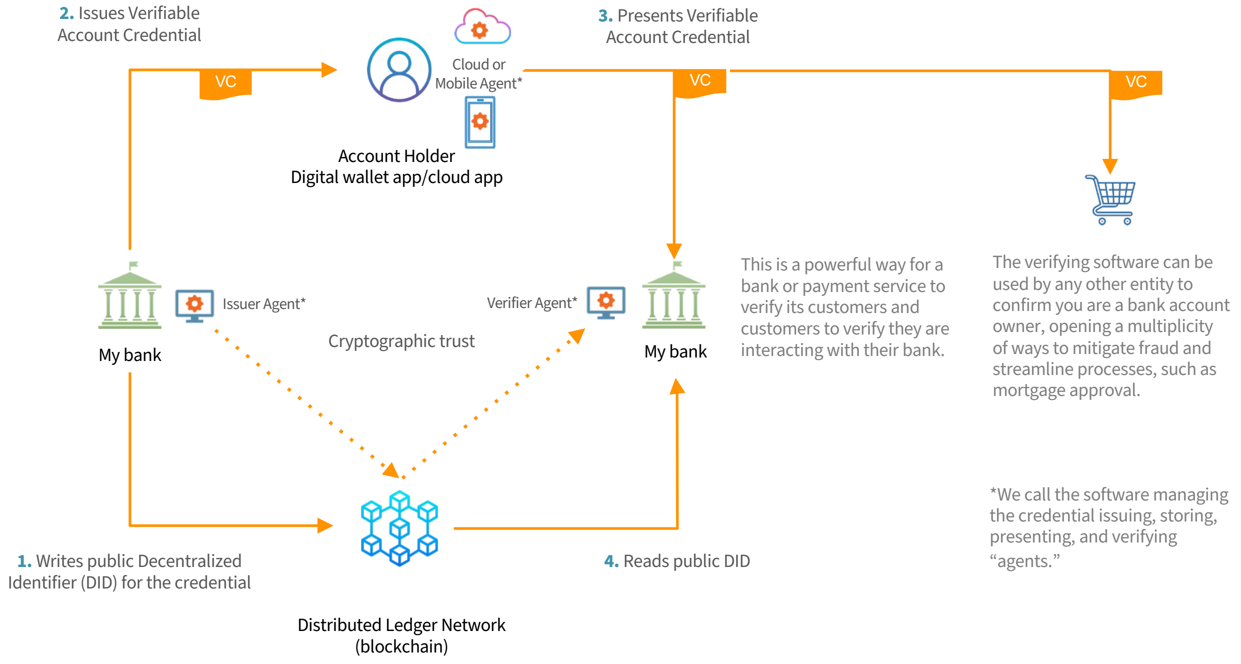
An example

Let's say I'm a bank. I've approved you for a new account by going through all the identity assurance steps needed to prove that you are who you say you are. I'm now going to offer you a way to prove you are an account holder by issuing you a verifiable credential for your account using the technology we've described.

Instead of a username or password, you'll use this credential to log into your account and for any account transactions. We'll know it's you logging into your account because we'll verify you are in control of your DID — and equally important, you'll know you are interacting with your bank and not someone pretending to be your bank.

I (the bank), send you a QR code either via email or a text message with the offer of an account credential. You scan the QR code with your mobile device. This connection, over DIDComm, enables you to accept and receive a verifiable credential, which you will store in special software on your mobile device — a digital wallet app — or in the cloud. This credential contains relevant details about your account.

Here's an important step: In creating the verifiable credential, the bank sends you your account details directly; they're in the verifiable credential. The bank will also write a cryptographic digital signature and metadata about the structure of the credential — called the schema — to the distributed ledger network. This information means the authenticity of the credential can be verified along with the integrity of the information contained within the credential, i.e., that the information hasn't been tampered with.


2. Issues Verifiable Account Credential
VC
Cloud or Mobile Agent*
Account Holder
Digital wallet app/cloud app
3. Presents Verifiable Account Credential
VC
VC
Issuer Agent*
My bank
Cryptographic trust
Verifier Agent*
My bank
1. Writes public Decentralized Identifier (DID) for the credential
Distributed Ledger Network (blockchain)
4. Reads public DID

This is a powerful way for a bank or payment service to verify its customers and customers to verify they are interacting with their bank.

The verifying software can be used by any other entity to confirm you are a bank account owner, opening a multiplicity of ways to mitigate fraud and streamline processes, such as mortgage approval.

*We call the software managing the credential issuing, storing, presenting, and verifying "agents."

A Trusted Digital Ecosystem

Let’s reiterate the iron law: *none of your personal information is stored on the distributed ledger*. The ledger only contains a list of the data fields contained in the credential, and your personal information that fills those fields is only on your device and in your control. If your personal data was written to the ledger, anyone could see it.

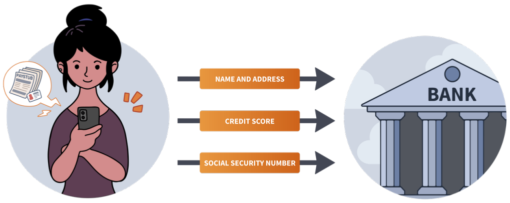
What if I lose my phone or someone steals it? Biometric binding and liveness checks ensure that it is you who is using your phone and digital wallet app. Backups can be held in the cloud. And verifiable credentials can be revoked and quickly reissued.

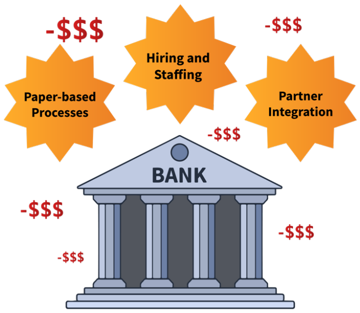
We describe this implementation of verifiable credentials as “Trusted Digital Ecosystems.” This is because each party is able to trust the source of information and the integrity of that information. And once they can do that, the ecosystem can expand to include new verifiers and issuers and use cases.

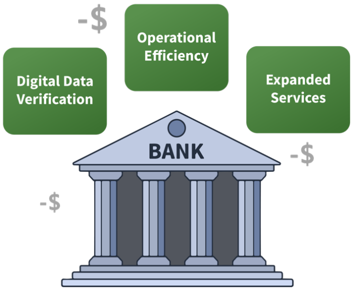
As an example, consider how verifiable credentials manage something like a mortgage application: Instead of having to gather numerous paper documents that either must be hand delivered or scanned and emailed, often repeatedly, you will be able to hold all this information through a variety of verifiable credentials — one for your bank statements, one for your pay stubs, one for your personal ID — and then combine them all to send to a broker.

These are the kinds of seamless processes that verifiable data enable.

PAYSTUBS
PHOTO ID
GIFT LETTERS
PROOF OF ADDRESS
BANK STATEMENTS
TAX RETURNS
RENTING HISTORY
CREDIT REPORTS
BANK







Integration and privacy

Make the technology that does this easy to integrate into existing infrastructure, easy to use by everyone, privacy-preserving, frictionless, scalable, and low cost!

There are several additional important points to make: We've already eliminated the need for third-party centralized databases to store information. But we can also share the data in a verifiable credential in ways that give us additional privacy protection. We can choose to selectively share specific data, based on the attribute field in the schema.

We can also prove data about ourselves in ways that don't require sharing the specific data. This is called a "zero-knowledge proof," an example of which would be the ability to answer "yes" or "no" to the question "are you over 18?"

Derivative credentials are another way of proving something without sharing specifics. Say you must pass a health test to enter a country. This data can be issued by a healthcare provider in the form of a verifiable credential. The credential can then be verified by a government authority, which in turn issues a second credential that only attests to the test results having been verified. The person could then use this credential to prove their health without having to share any health or personal data whatsoever.

Let’s summarize

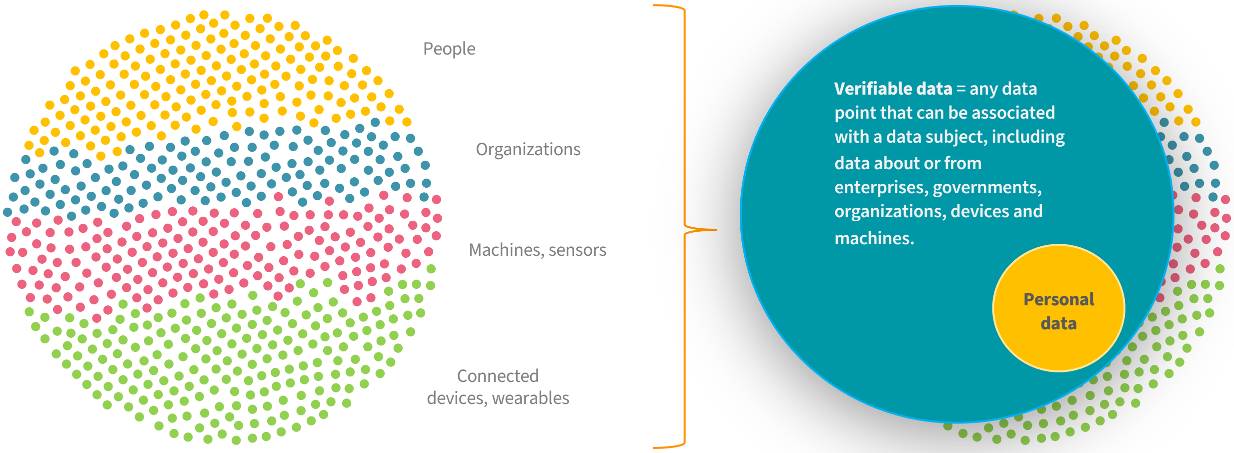
- I can create a digital identity-identifier.
- I can prove that I am in control of my digital identity-identifier.
- I can accept personal data in the form of a verifiable credential and can prove the authenticity of the credential and the integrity of the data.
- I can control who I share this data with through authenticated, encrypted communication channels.
- And anyone with verifying software can verify a credential.

Here’s the thing that’s really powerful: we’re not just talking about the kind of personal data contained in ID cards, important thought that is; any information that can be conveyed through a credential schema is now verifiable along with the source of that information.

Think about how this technology will manage something like a mortgage application: Instead of having to gather a plethora of paper documentation that either has to be hand delivered or scanned and emailed, usually over and over again, you will be able to hold all this information through a variety of verifiable credentials — one for your bank statements, one for your pay stubs, one for your personal ID — and then combine them all to send to a broker. These are the kinds of seamless processes that verifiable data enable.

In sum, with this technology we have a solution to the internet’s missing identity layer for people and organizations and things, and we have a way to prove the authenticity of the data shared by people, organizations, and things.

How web 2.0 and web3 will co-evolve


People
Organizations
Machines, sensors
Connected devices, wearables
Verifiable data = any data point that can be associated with a data subject, including data about or from enterprises, governments, organizations, devices and machines.
Personal data

Unlimited digital identities
holding their data off chain in verifiable credentials

Implementation

Decentralized identity and verifiable data are not wishful thinking technology or future ware waiting to be developed. Right now, this technology is being implemented across many different sectors from financial services to supply chain management to device monitoring and to travel. The ability to create and use verifiable data will eventually transform everything, not least because the web cannot evolve without a reliable way to manage identity and verification: the cost in fraud and friction from existing systems is untenable while the demand for seamless digital interaction is voracious.

One of the keys to this digital transformation is that verifiable credential technology can be layered onto existing systems without these systems having to be replaced. This avoids the high infrastructural costs and endless timelines of shifting to new systems.

Governance can be automated to make complex information flows easy and offline verification possible. Interoperability, in terms of shared protocols and standards, means that credentials will be usable across industries and countries.

Finally, open-source codebases mean that the technology can evolve in a sustainable way while being accessible to anyone (open source means the code for the technology is open to everyone to use and develop on). Open-source technologies drive a virtuous cycle of innovation: as the number of implementations increase, the developer community grows, which in turn drives more adoption and innovation. The codebase also becomes more robust as it is used in multiple contexts by many parties. Open source is now a standard requirement in public-sector decentralized identity projects.

In terms of user experience, the technology involves nothing more than downloading an app and swiping. For those without mobile devices, there are cloud-based equivalents.

Hopefully, this beginner's guide gives you a sense of what this technology does, how it works, and how powerful and useful it is. By necessity, it leaves out many technical terms and subtleties, but there are many resources available for you to dive deeper—not least The Indicio Academy or Proven Sandbox or the Hyperledger Foundation. As with many new technologies, the best way to learn about decentralized identity is to actually use it.

Verifiable credential technology is evolving rapidly: The photo above shows Indicio in the process of turning a physical passport into a verifiable credential as part of its work with SITA, the world's leading supplier of IT to the air transport sector. The goal is a Digital Travel Credential that will allow seamless check in and border crossing. The technology was successfully trialed in March 2023, when air passengers from around the world used a DTC to enter Aruba. For more details, here's a link to a DTC video and a link to press conference.

The time to innovate is now!

Indicio's decentralized identity products accelerate time to value and reduce the operational cost and complexity of building and delivering verifiable credentials include:

- Indicio Sandbox: create pilot projects, demonstrations, in a complete hosted environment while you learn how the technology works.
- Holdr+: collect verifiable credentials and communicate with connections using a widely-available mobile application.
- Indicio Academy: learn how to build and deploy successful systems in the only training and certification program, specific to the industry.
- Indicio Network: rely on our professionally managed, globally distributed ledger network that underpins our products.

The core of Indicio is Indicio Proven™, which provides everything you need to issue, hold, share, and verify digital information at scale, and available as turnkey software, hosted either on one of Indicio's cloud partners, or the hosting service of your choice.

You also can explore industry-specific products: Indicio Proven Works, a turnkey solution for employee credentialing and Indicio Proven Finance that makes it easy to issue financial documents as verifiable credentials in the financial services industry.

Confidently build and scale open-source verifiable credentials with a streamlined experience across your existing systems, services, and applications while we manage the rest!

Contact us today:
Sales@Indicio.tech
Join our mailing list
Get a free demo

Connect with us:

- Newsletter
- Blog
- YouTube
- LinkedIn
- Twitter